

IN THE CLAIMS

Please cancel claims 26, 47 and 66, and amend claims 24, 44, 45 and 53 as follows:

1-23. (CANCELED)

24. (CURRENTLY AMENDED) A computer-implemented system for performing data mining applications, comprising:

(a) a computer having one or more data storage devices connected thereto, wherein a relational database is stored on one or more of the data storage devices;

(b) a relational database management system, executed by the computer, for accessing the relational database stored on the data storage devices by executing Structured Query Language (SQL) statements;

(c) an analytic application programming interface (API), executed by the computer, for invoking one or more scalable data mining functions comprised of SQL statements for execution by the relational database management system, wherein the scalable data mining functions identify and interpret patterns in the relational database; and

(d) one or more analytic algorithms, executed by the computer, for interfacing to the analytic API to invoke the scalable data mining functions, wherein the scalable data mining functions process data collections stored in the relational database and produce results that are stored in the relational database.

25. (PREVIOUSLY PRESENTED) The system of claim 24 above, wherein the computer comprises a parallel processing computer comprised of a plurality of nodes, and each node executes one or more threads of the relational database management system to provide parallelism in the data mining operations.

26. (CANCELED)

27. (PREVIOUSLY PRESENTED) The system of claim 24, wherein the scalable data mining functions are created by parameterizing and instantiating the analytic API.

28. (PREVIOUSLY PRESENTED) The system of claim 24, wherein the scalable data mining functions are dynamically generated SQL statements comprised of combined phrases with substituting values therein based on parameters supplied to the analytic API.

29. (CANCELED)

30. (PREVIOUSLY PRESENTED) The system of claim 24, wherein the scalable data mining functions comprise Data Description functions that provide descriptive statistical functions.

31. (PREVIOUSLY PRESENTED) The system of claim 30, wherein the Data Description functions comprise:

(1) descriptive statistics for one or more numeric columns, wherein the statistics are selected from a group comprising count, minimum, maximum, mean, standard deviation, standard mean error, variance, coefficient of variance, skewness, kurtosis, uncorrected sum of squares, corrected sum of squares, and quantiles,

(2) a count of values for a column,

(3) a calculated modality for a column,

(4) one or more bin numeric columns of counts with overlay and statistics options,

(5) one or more automatically sub-binned numeric columns giving additional counts and isolated frequently occurring individual values

(6) a computed frequency of one or more column values,

(7) a computed frequency of values for pairs of columns in a column list,

(8) a Pearson Product-Moment Correlation matrix,

(9) a Covariance matrix,

(10) a sum of squares and cross-products matrix, or

(11) a count of overlapping column values in one or more combinations of tables.

32. (PREVIOUSLY PRESENTED) The system of claim 24, wherein the scalable data mining functions comprise Data Derivation functions that provide column derivations or transformations.

33. (PREVIOUSLY PRESENTED) The system of claim 32, wherein the Data Derivation functions comprise:

- (1) a derived binned numeric column wherein a new column is bin number,
- (2) a n-valued categorical column dummy-coded into "n" 0/1 values,
- (3) a n-valued categorical column recoded into n or less new values,
- (4) one or more numeric columns scaled via range transformation,
- (5) one or more columns scaled to a z-score that is a number of standard deviations

from a mean,

- (6) one or more numeric columns scaled via a sigmoidal transformation function,
- (7) one or more numeric columns scaled via a base 10 logarithm function,
- (8) one or more numeric columns scaled via a natural logarithm function,
- (9) one or more numeric columns scaled via an exponential function,
- (10) one or more numeric columns raised to a specified power,
- (11) one or more numeric columns derived via user defined transformation function,
- (12) one or more new columns derived by ranking one or more columns or expressions

based on order,

- (13) one or more new columns derived with quantile 0 to n-1 based on order and n,
- (14) a cumulative sum of a value expression based on a sort expression,
- (15) a moving average of a value expression based on a width and order,
- (16) a moving sum of a value expression based on a width and order,
- (17) a moving difference of a value expression based on a width and order,
- (18) a moving linear regression value derived from an expression, width, and order,
- (19) a multiple account/product ownership bitmap,
- (20) a product ownership bitmap over multiple time periods,
- (21) one or more counts, amount, percentage means and intensities derived from a

transaction summary,

- (22) one or more variabilities derived from transaction summary data,
- (23) one or more derived trigonometric values and their inverses, including sin, arcsin,

cos, arccos, csc, arccsc, sec, arcsec, tan, arctan, cot, and arccot, or

(24) one or more derived hyperbolic values and their inverses, including sinh, arcsinh, cosh, arccosh, csch, arccsch, sech, arcsech, tanh, arctanh, coth, and arccoth.

34. (PREVIOUSLY PRESENTED) The system of claim 24, wherein the scalable data mining functions comprise Data Reduction functions that provide matrix building operations to reduce the amount of data required for analytic algorithms.

35. (PREVIOUSLY PRESENTED) The system of claim 34, wherein the Data Reduction functions comprise:

- (1) build one or more data reduction matrices from a group comprising: (i) a Pearson-Product Moment Correlations matrix; (ii) a Covariances matrix; and (iii) a Sum of Squares and Cross Products (SSCP) matrix,
- (2) export a resultant matrix, or
- (3) restart a matrix operation.

36. (PREVIOUSLY PRESENTED) The system of claim 24, wherein the scalable data mining functions comprise Data Reorganization functions that provide an ability to reorganize data by joining or de-normalizing pre-processed results into a wide analytic data set.

37. (PREVIOUSLY PRESENTED) The system of claim 36, wherein the Data Reorganization functions comprise:

- (1) create a de-normalized new table by removing one or more key columns, or
- (2) join a plurality of tables or views into a combined result table.

38. (PREVIOUSLY PRESENTED) The system of claim 24, wherein the scalable data mining functions comprise a Data Sampling function that provides an ability to construct a new table containing a randomly selected subset of the rows in an existing table or view.

39. (PREVIOUSLY PRESENTED) The system of claim 38, wherein the Data Sampling function selects one or more data samples of specified sizes from a table.

40. (PREVIOUSLY PRESENTED) The system of claim 24, wherein the scalable data mining functions comprise a Data Partitioning function that provides an ability to construct a new table containing at least one randomly selected subset of the rows in an existing table or view, wherein the subsets are mutually distinct but all-inclusive subsets of data.

41. (PREVIOUSLY PRESENTED) The system of claim 40, wherein the Data Partitioning function selects one or more data partitions from a table using a database internal hashing technique.

42. (PREVIOUSLY PRESENTED) The system of claim 24, wherein results of the data mining operations are stored in the relational database.

43. (PREVIOUSLY PRESENTED) The system of claim 24, wherein the relational database management system further comprises an analytical logical data model that stores metadata and processing results from the Scalable Data Mining Functions.

44. (CURRENTLY AMENDED) A computer-implemented method for performing data mining applications, comprising:

(a) storing a relational database on one or more data storage devices connected to a computer;

(b) accessing the relational database stored on the data storage devices using a relational database management system by executing Structured Query Language (SQL) statements on the computer;

(c) executing an analytic application programming interface (API), on the computer, for invoking one or more scalable data mining functions comprised of SQL statements for execution by the relational database management system, wherein the scalable data mining functions identify and interpret patterns in the relational database; and

(d) executing one or more analytic algorithms, on the computer, for interfacing to the analytic API to invoke the scalable data mining functions, wherein the scalable data mining functions process data collections stored in the relational database and produce results that are stored in the relational database.

45. (CURRENTLY AMENDED) An article of manufacture tangibly embodying logic for performing a computer-implemented method for performing data mining applications, comprising:

(a) storing a relational database on one or more data storage devices connected to a computer;

(b) accessing the relational database stored on the data storage devices using a relational database management system by executing Structured Query Language (SQL) statements on the computer;

(c) executing an analytic application programming interface (API), on the computer, for invoking one or more scalable data mining functions comprised of SQL statements for execution by the relational database management system, wherein the scalable data mining functions identify and interpret patterns in the relational database; and

(d) executing one or more analytic algorithms, on the computer, for interfacing to the analytic API to invoke the scalable data mining functions, wherein the scalable data mining functions process data collections stored in the relational database and produce results that are stored in the relational database.

46. (PREVIOUSLY PRESENTED) The method of claim 44 above, wherein the computer comprises a parallel processing computer comprised of a plurality of nodes, and each node executes one or more threads of the relational database management system to provide parallelism in the data mining operations.

47. (CANCELED)

48. (PREVIOUSLY PRESENTED) The method of claim 44, wherein the scalable data mining functions are created by parameterizing and instantiating the analytic API.

49. (PREVIOUSLY PRESENTED) The method of claim 44, wherein the scalable data mining functions are dynamically generated SQL statements comprised of combined phrases with substituting values therein based on parameters supplied to the analytic API.

50. (CANCELED)

51. (PREVIOUSLY PRESENTED) The method of claim 44, wherein the scalable data mining functions comprise Data Description functions that provide descriptive statistical functions.

52. (PREVIOUSLY PRESENTED) The method of claim 51, wherein the Data Description functions comprise:

(1) descriptive statistics for one or more numeric columns, wherein the statistics are selected from a group comprising count, minimum, maximum, mean, standard deviation, standard mean error, variance, coefficient of variance, skewness, kurtosis, uncorrected sum of squares, corrected sum of squares, and quantiles,

(2) a count of values for a column,

(3) a calculated modality for a column,

(4) one or more bin numeric columns of counts with overlay and statistics options,

(5) one or more automatically sub-binned numeric columns giving additional counts and isolated frequently occurring individual values

(6) a computed frequency of one or more column values,

(7) a computed frequency of values for pairs of columns in a column list,

(8) a Pearson Product-Moment Correlation matrix,

(9) a Covariance matrix,

(10) a sum of squares and cross-products matrix, or

(11) a count of overlapping column values in one or more combinations of tables.

53. (PREVIOUSLY PRESENTED) The method of claim 51, wherein the scalable data mining functions comprise Data Derivation functions that provide column derivations or transformations.

54. (PREVIOUSLY PRESENTED) The method of claim 53, wherein the Data Derivation functions comprise:

- (1) a derived binned numeric column wherein a new column is bin number,
- (2) a n-valued categorical column dummy-coded into "n" 0/1 values,
- (3) a n-valued categorical column recoded into n or less new values,
- (4) one or more numeric columns scaled via range transformation,
- (5) one or more columns scaled to a z-score that is a number of standard deviations from a mean,
- (6) one or more numeric columns scaled via a sigmoidal transformation function,
- (7) one or more numeric columns scaled via a base 10 logarithm function,
- (8) one or more numeric columns scaled via a natural logarithm function,
- (9) one or more numeric columns scaled via an exponential function,
- (10) one or more numeric columns raised to a specified power,
- (11) one or more numeric columns derived via user defined transformation function,
- (12) one or more new columns derived by ranking one or more columns or expressions based on order,
- (13) one or more new columns derived with quantile 0 to n-1 based on order and n,
- (14) a cumulative sum of a value expression based on a sort expression,
- (15) a moving average of a value expression based on a width and order,
- (16) a moving sum of a value expression based on a width and order,
- (17) a moving difference of a value expression based on a width and order,
- (18) a moving linear regression value derived from an expression, width, and order,
- (19) a multiple account/product ownership bitmap,
- (20) a product ownership bitmap over multiple time periods,
- (21) one or more counts, amount, percentage means and intensities derived from a transaction summary,
- (22) one or more variabilities derived from transaction summary data,
- (23) one or more derived trigonometric values and their inverses, including sin, arcsin, cos, arccos, csc, arccsc, sec, arcsec, tan, arctan, cot, and arccot, or
- (24) one or more derived hyperbolic values and their inverses, including sinh, arcsinh, cosh, arccosh, csch, arccsch, sech, arcsech, tanh, arctanh, coth, and arccoth.



55. (PREVIOUSLY PRESENTED) The method of claim 44, wherein the scalable data mining functions comprise Data Reduction functions that provide matrix building operations to reduce the amount of data required for analytic algorithms.

56. (PREVIOUSLY PRESENTED) The method of claim 55, wherein the Data Reduction functions comprise:

- (1) build one or more data reduction matrices from a group comprising: (i) a Pearson-Product Moment Correlations matrix; (ii) a Covariances matrix; and (iii) a Sum of Squares and Cross Products (SSCP) matrix,
- (2) export a resultant matrix, or
- (3) restart a matrix operation.

57. (PREVIOUSLY PRESENTED) The method of claim 44, wherein the scalable data mining functions comprise Data Reorganization functions that provide an ability to reorganize data by joining or de-normalizing pre-processed results into a wide analytic data set.

58. (PREVIOUSLY PRESENTED) The method of claim 57, wherein the Data Reorganization functions comprise:

- (1) create a de-normalized new table by removing one or more key columns, or
- (2) join a plurality of tables or views into a combined result table.

59. (PREVIOUSLY PRESENTED) The method of claim 44, wherein the scalable data mining functions comprise a Data Sampling function that provides an ability to construct a new table containing a randomly selected subset of the rows in an existing table or view.

60. (PREVIOUSLY PRESENTED) The method of claim 59, wherein the Data Sampling function selects one or more data samples of specified sizes from a table.

61. (PREVIOUSLY PRESENTED) The method of claim 44, wherein the scalable data mining functions comprise a Data Partitioning function that provides an ability to construct a

new table containing at least one randomly selected subset of the rows in an existing table or view, wherein the subsets are mutually distinct but all-inclusive subsets of data.

62. (PREVIOUSLY PRESENTED) The method of claim 61, wherein the Data Partitioning function selects one or more data partitions from a table using a database internal hashing technique.

63. (PREVIOUSLY PRESENTED) The method of claim 44, wherein results of the data mining operations are stored in the relational databases.

64. (PREVIOUSLY PRESENTED) The method of claim 44, wherein the relational database management system further comprises an analytical logical data model that stores metadata and processing results from the Scalable Data Mining Functions.

65. (PREVIOUSLY PRESENTED) The article of claim 45 above, wherein the computer comprises a parallel processing computer comprised of a plurality of nodes, and each node executes one or more threads of the relational database management system to provide parallelism in the data mining operations.

66. (CANCELED)

67. (PREVIOUSLY PRESENTED) The article of claim 45, wherein the scalable data mining functions are created by parameterizing and instantiating the analytic API.

68. (PREVIOUSLY PRESENTED) The article of claim 45, wherein the scalable data mining functions are dynamically generated SQL statements comprised of combined phrases with substituting values therein based on parameters supplied to the analytic API.

69. (CANCELED)

70. (PREVIOUSLY PRESENTED) The article of claim 45, wherein the scalable data mining functions comprise Data Description functions that comprise descriptive statistical functions.

71. (PREVIOUSLY PRESENTED) The article of claim 70, wherein the Data Description functions comprise:

(1) descriptive statistics for one or more numeric columns, wherein the statistics are selected from a group comprising count, minimum, maximum, mean, standard deviation, standard mean error, variance, coefficient of variance, skewness, kurtosis, uncorrected sum of squares, corrected sum of squares, and quantiles,

(2) a count of values for a column,

(3) a calculated modality for a column,

(4) one or more bin numeric columns of counts with overlay and statistics options,

(5) one or more automatically sub-binned numeric columns giving additional counts and isolated frequently occurring individual values

(6) a computed frequency of one or more column values,

(7) a computed frequency of values for pairs of columns in a column list,

(8) a Pearson Product-Moment Correlation matrix,

(9) a Covariance matrix,

(10) a sum of squares and cross-products matrix, or

(11) a count of overlapping column values in one or more combinations of tables.

72. (PREVIOUSLY PRESENTED) The article of claim 45, wherein the scalable data mining functions comprise Data Derivation functions that provide column derivations or transformations.

73. (PREVIOUSLY PRESENTED) The article of claim 72, wherein the Data Derivation functions comprise:

(1) a derived binned numeric column wherein a new column is bin number,

(2) a n-valued categorical column dummy-coded into "n" 0/1 values,

(3) a n-valued categorical column recoded into n or less new values,

- (4) one or more numeric columns scaled via range transformation,
- (5) one or more columns scaled to a z-score that is a number of standard deviations from a mean,
- (6) one or more numeric columns scaled via a sigmoidal transformation function,
- (7) one or more numeric columns scaled via a base 10 logarithm function,
- (8) one or more numeric columns scaled via a natural logarithm function,
- (9) one or more numeric columns scaled via an exponential function,
- (10) one or more numeric columns raised to a specified power,
- (11) one or more numeric columns derived via user defined transformation function,
- (12) one or more new columns derived by ranking one or more columns or expressions based on order,
- (13) one or more new columns derived with quantile 0 to n-1 based on order and n,
- (14) a cumulative sum of a value expression based on a sort expression,
- (15) a moving average of a value expression based on a width and order,
- (16) a moving sum of a value expression based on a width and order,
- (17) a moving difference of a value expression based on a width and order,
- (18) a moving linear regression value derived from an expression, width, and order,
- (19) a multiple account/product ownership bitmap,
- (20) a product ownership bitmap over multiple time periods,
- (21) one or more counts, amount, percentage means and intensities derived from a transaction summary,
- (22) one or more variabilities derived from transaction summary data,
- (23) one or more derived trigonometric values and their inverses, including sin, arcsin, cos, arccos, csc, arccsc, sec, arcsec, tan, arctan, cot, and arccot, or
- (24) one or more derived hyperbolic values and their inverses, including sinh, arcsinh, cosh, arccosh, csch, arccsch, sech, arcsech, tanh, arctanh, coth, and arccoth.

74. (PREVIOUSLY PRESENTED) The article of claim 45, wherein the scalable data mining functions comprise Data Reduction functions that provide matrix building operations to reduce the amount of data required for analytic algorithms.

75. (PREVIOUSLY PRESENTED) The article of claim 74, wherein the Data Reduction functions comprise:

- (1) build one or more data reduction matrices from a group comprising: (i) a Pearson-Product Moment Correlations matrix; (ii) a Covariances matrix; and (iii) a Sum of Squares and Cross Products (SSCP) matrix,
- (2) export a resultant matrix, and or
- (3) restart a matrix operation.

76. (PREVIOUSLY PRESENTED) The article of claim 45, wherein the scalable data mining functions comprise Data Reorganization functions that provide an ability to reorganize data by joining or de-normalizing pre-processed results into a wide analytic data set.

77. (PREVIOUSLY PRESENTED) The article of claim 76, wherein the Data Reorganization functions comprise:

- (1) create a de-normalized new table by removing one or more key columns, or
- (2) join a plurality of tables or views into a combined result table.

78. (PREVIOUSLY PRESENTED) The article of claim 45, wherein the scalable data mining functions comprise a Data Sampling function that provides an ability to construct a new table containing a randomly selected subset of the rows in an existing table or view.

79. (PREVIOUSLY PRESENTED) The article of claim 78, wherein the Data Sampling function selects one or more data samples of specified sizes from a table.

80. (PREVIOUSLY PRESENTED) The article of claim 45, wherein the scalable data mining functions comprise a Data Partitioning function that provides an ability to construct a new table containing at least one randomly selected subset of the rows in an existing table or view, wherein the subsets are mutually distinct but all-inclusive subsets of data.

81. (PREVIOUSLY PRESENTED) The article of claim 80, wherein the Data Partitioning function selects one or more data partitions from a table using a database internal hashing technique.

82. (PREVIOUSLY PRESENTED) The article of claim 45, wherein results of the data mining operations are stored in the relational database.

83. (PREVIOUSLY PRESENTED) The article of claim 45, wherein the relational database management system further comprises an analytical logical data model that stores metadata and processing results from the Scalable Data Mining Functions.